Analysis on Aviation Incidents and Crashes

Presented by Abhay Rawat

➢ Problem Statement :-

- In this project, I am going to analyze the data of Aviation Crashes and Incidents that occurred since 1908 to 2008 because it provides a comprehensive historical record of aviation safety incidents over a century . It includes information on the types of planes, the number of people aboard, fatalities, routes, operators, and summaries.
- The objective of this project is to understand historical accident trends, identify key areas for improvement, and inform future strategies to enhance aviation safety and prevent accidents.

WORKFLOW:-

- I divide my work flow into 3 steps:

- EDA is divided into 3 Analysis:-

1) **Univariate Analysis:** Here, I am analysing only one variable.
2) **Bivariate Analysis:** Here, I am analysing only two variable.
3) **Multivariate Analysis:** Here, I am analysing multivariable.

Data Collection and Understanding:-

- Data Collection and Understanding are the most important part of Exploratory Data and Analysis(EDA).So, I have data of Aviation Crashes and Incidents that I have to explore. This data contains 5268 rows and 14 columns. Let's understand the columns.

- **Data Description:**
  - Index: It contains Index of the Rows
  - Date: It contains Date of the Incidents
  - Time: It contains Time of the Incidents
  - Location: It contains the Location of the Crashes and Incidents
  - Operator: It contains the Operator name of the Aircraft
  - Flight : It contains Flight Number
  - Route : It contains Route of the Aircraft in the Incidents
  - Type: This columns contains the information about the Type of the Plane
  - Registration: It contains the Registration number of the aircraft
  - Cn/In: It contains the construction number and serial number of the aircraft
  - Aboard: This columns contains the information about the number of people on board the aircraft
  - Fatalities: It contains the information about the number of the the fatalities in the incidents
  - Ground: It contains the information about the number of people on the ground killed in the incidents
  - Summary: This columns contains the information about the summary of the incidents

Data Cleaning and Manipulation

- The given data has 5268 rows and 14 columns. We can view its shape using the '.shape' function.
- I have dropped unwanted columns such as "Index , Flight #, Registration , Cn/In".
- Missing value in the "Time", "Route", and "Summary" columns were filled with mean, backward fill, mode values using the '.fillna() ' function , and the remaining missing values were deleted.
- Data types of "Date" and "Time" columns have been changed form object to date time format.
- New columns like "Year", "Crash Region", and "Total Fatalities" were created.
- The year was extracted from the date columns and placed in the new "Year" columns.
- The location of the crash incident was extracted from the Location column, and a new column called "Total Fatalities" was created by summing the "Fatalities" and " Ground Fatalities " columns.

EDA: Univariate Analysis

Figure - 1

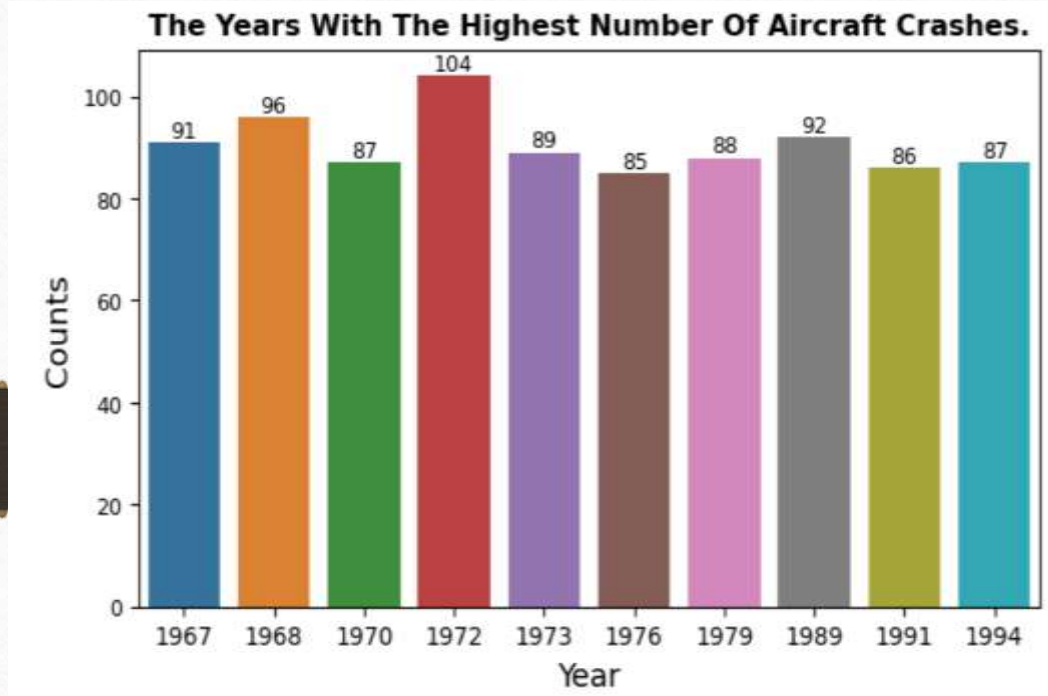
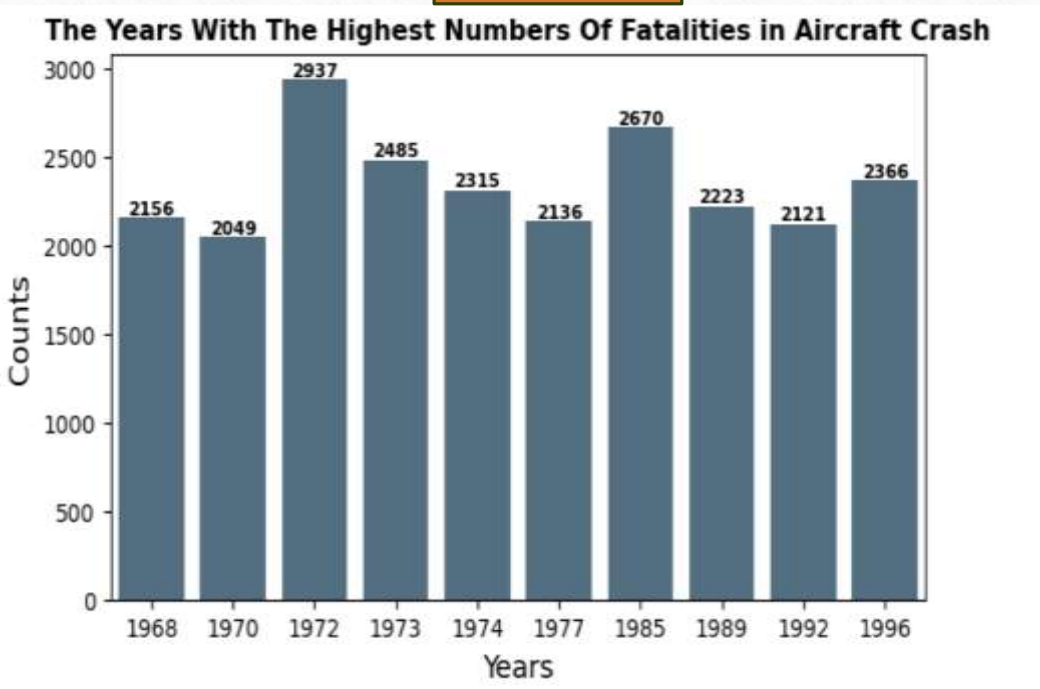


Figure - 2



Conclusions:

1-The maximum number of plane crashes occurred in 1972 with 104 crashes, followed by 1968 with 96 crashes.

2-The maximum number of fatalities occurred in 1972 with 2,937 fatalities, followed by 1985 with 2,670 fatalities.

EDA: Bivariate Analysis and Multivariate

Figure - 3

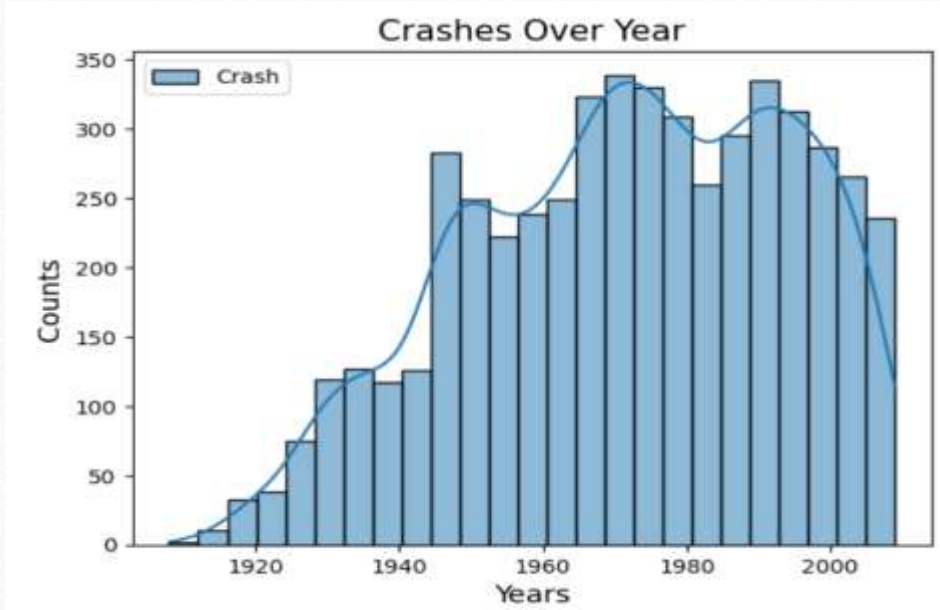
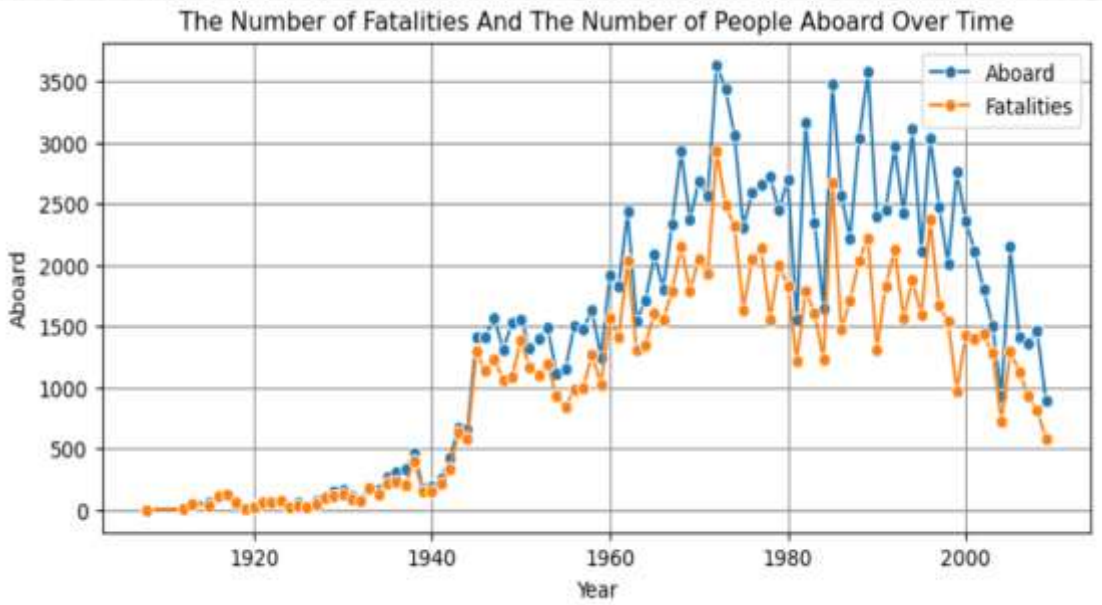


Figure - 4



Conclusions:

1- As shown in Figure 1, most of the incidents occurred between 1942 and 2000.

2- In Figure 2, we can see that the number of fatalities and the number of people aboard were highest between 1940 and 2000, after which they declined.

Figure - 5

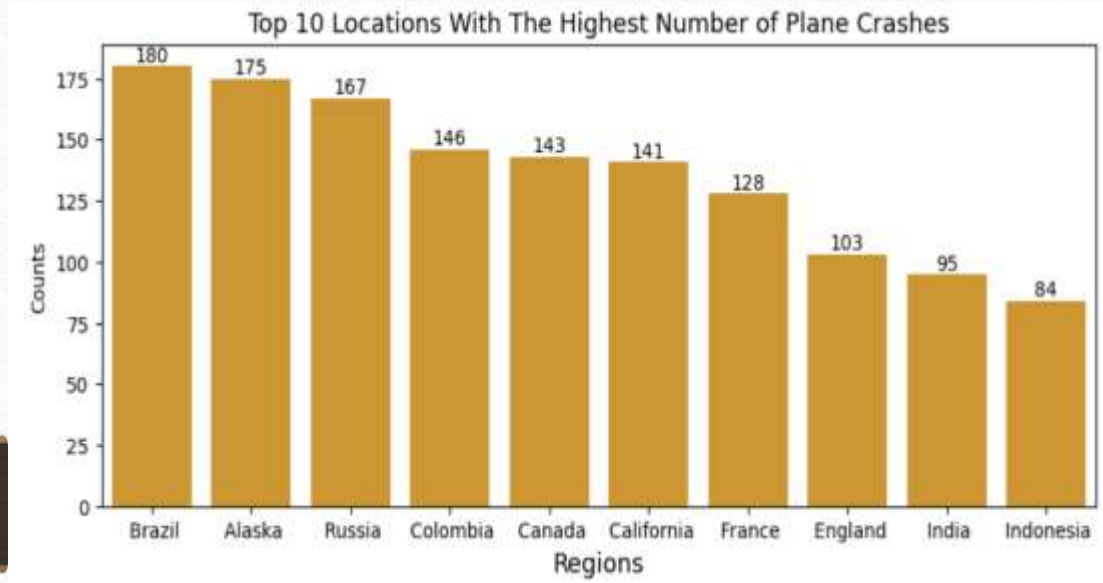
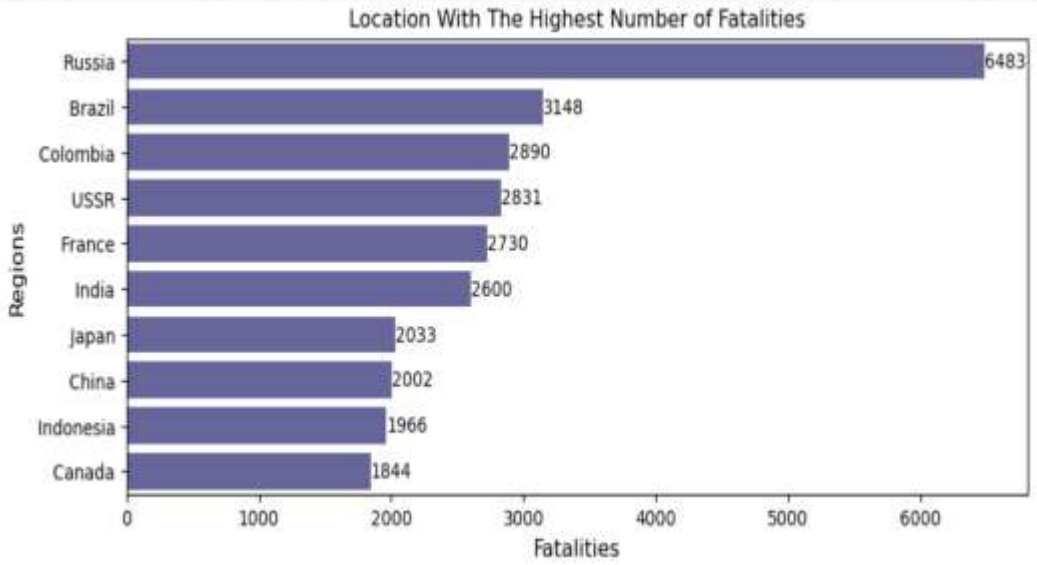


Figure - 6



Conclusions:

1- As we can see in the figure -1,Brazil has the highest number of plane crashes with 180 incidents, followed by Alaska with 175 and Russia with 167

2- In figure-2,Russia had the highest number of fatalities with 6483, followed by Brazil with 3148 and Colombia with 2890.

3- As we can see, most plane crashes occurred during takeoff. The reasons for these crashes may include bad weather, fogs, enigine failure, technical errors,collide with mountains,old planes,poor runway,Hijack,war, or pilot's mistakes etc.

Figure-7

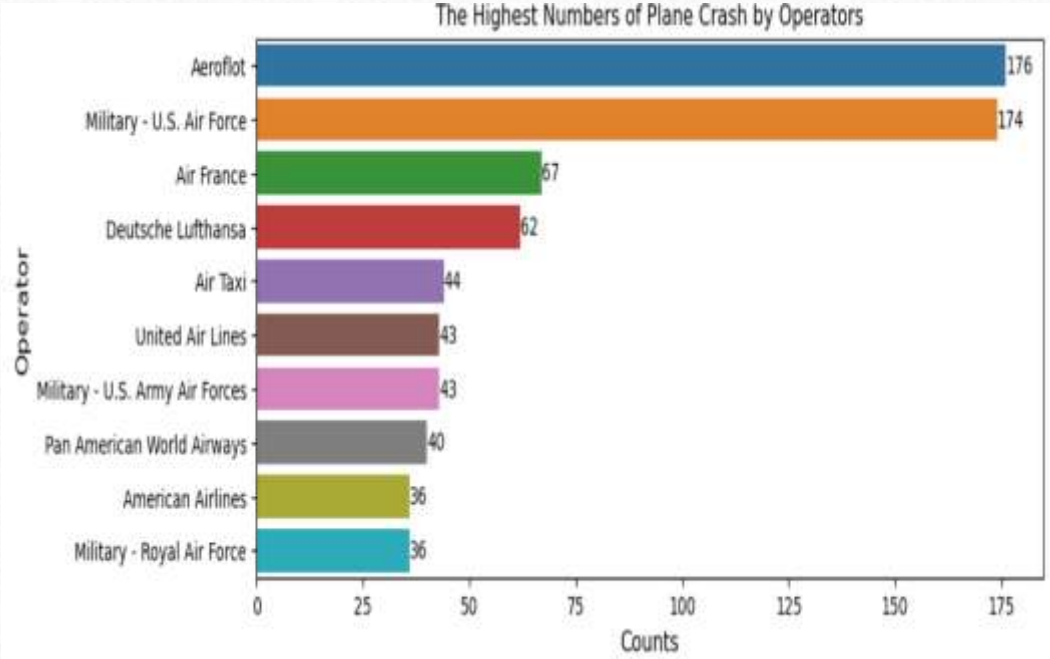
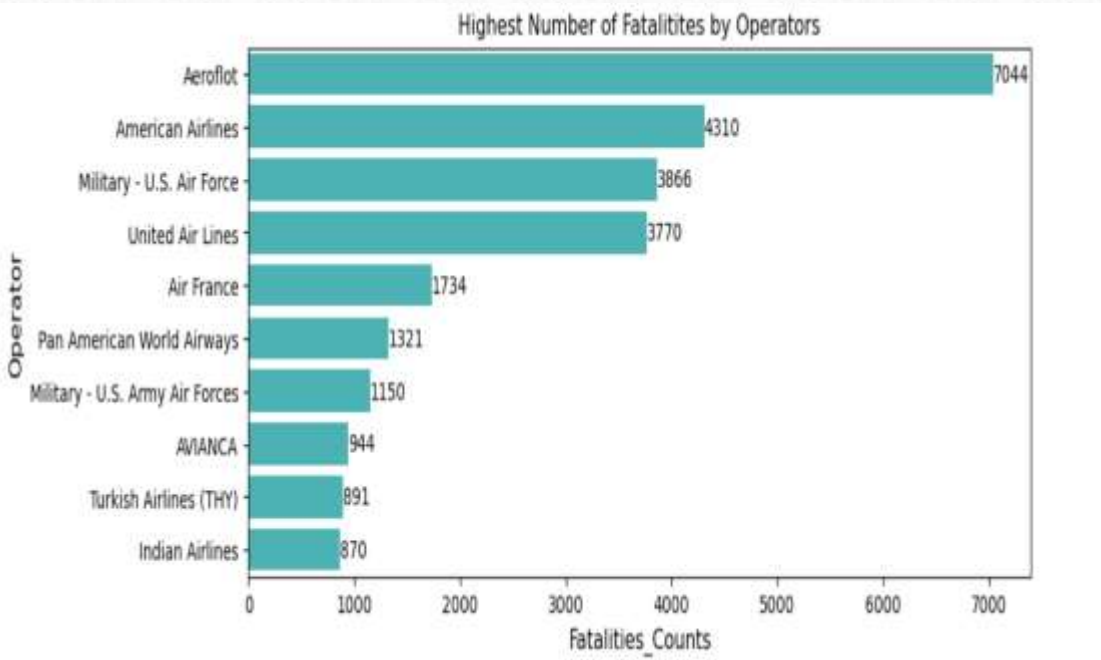


Figure-8



Conclusions:

1- As we can see in the figure -1, We can see in the graph above that 'Aeroflot' had 176 crashes, followed by the Military-U.S. Air Force with 174.

2- In figure-2,We can see in the graph above that 'Aeroflot' had 7044 crashes, followed by American Airlines with 4310 and the Military-U.S. Air Force with 3770.

Figure-9

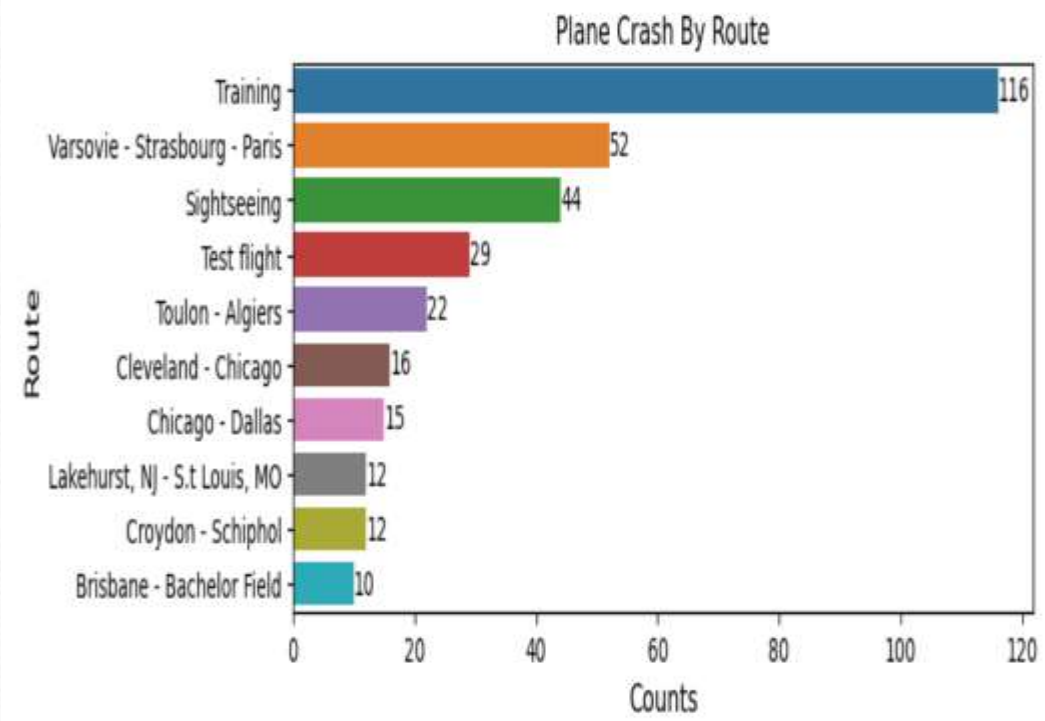
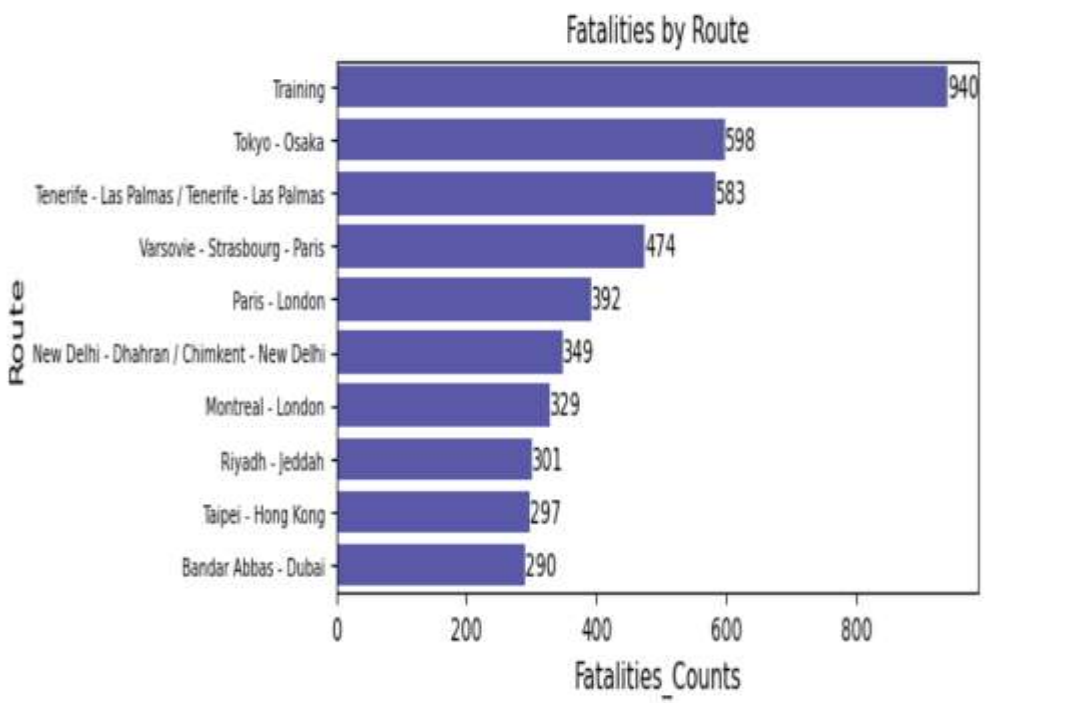


Figure-10



Conclusions:

1- In the above figure, the highest number of plane crashes occurred on the Training route, with 116 crashes.

2- In the above figure, the highest number of fatalities occurred on the Training route, with 940 fatalities, followed by the Tokyo-Osaka route with 598 fatalities.

Figure-11

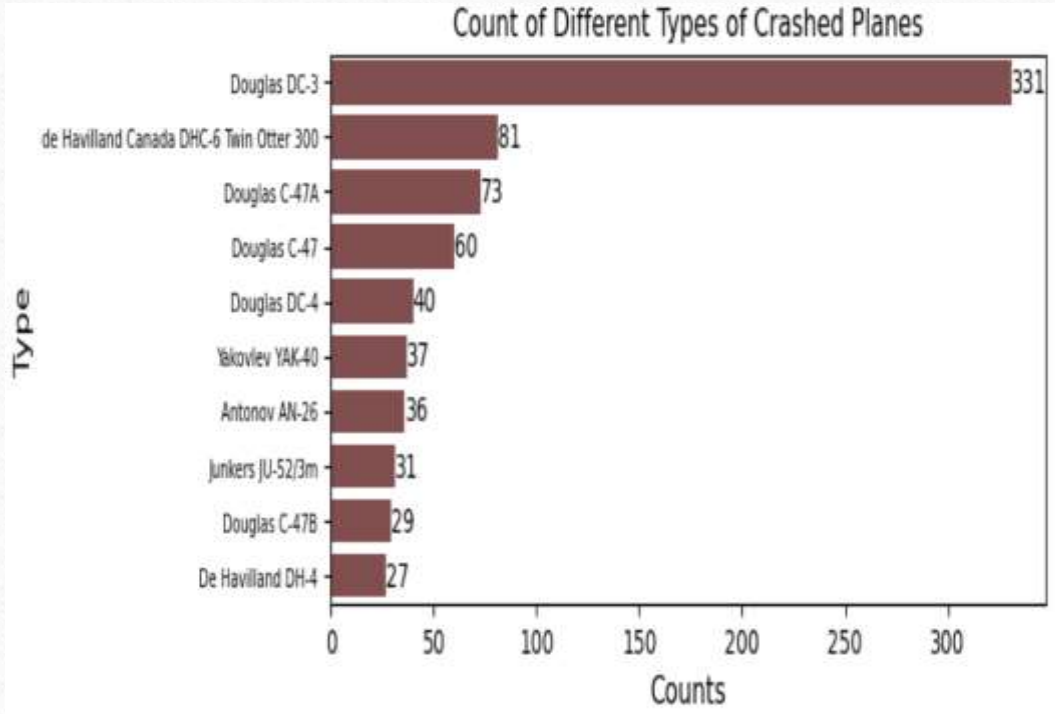
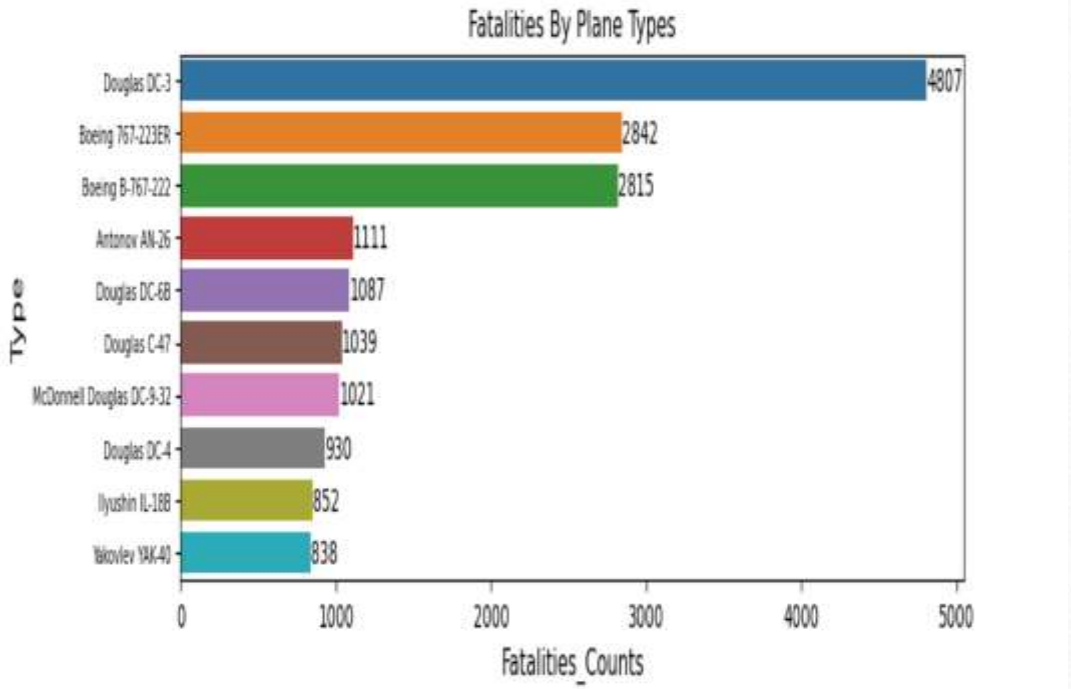


Figure - 12

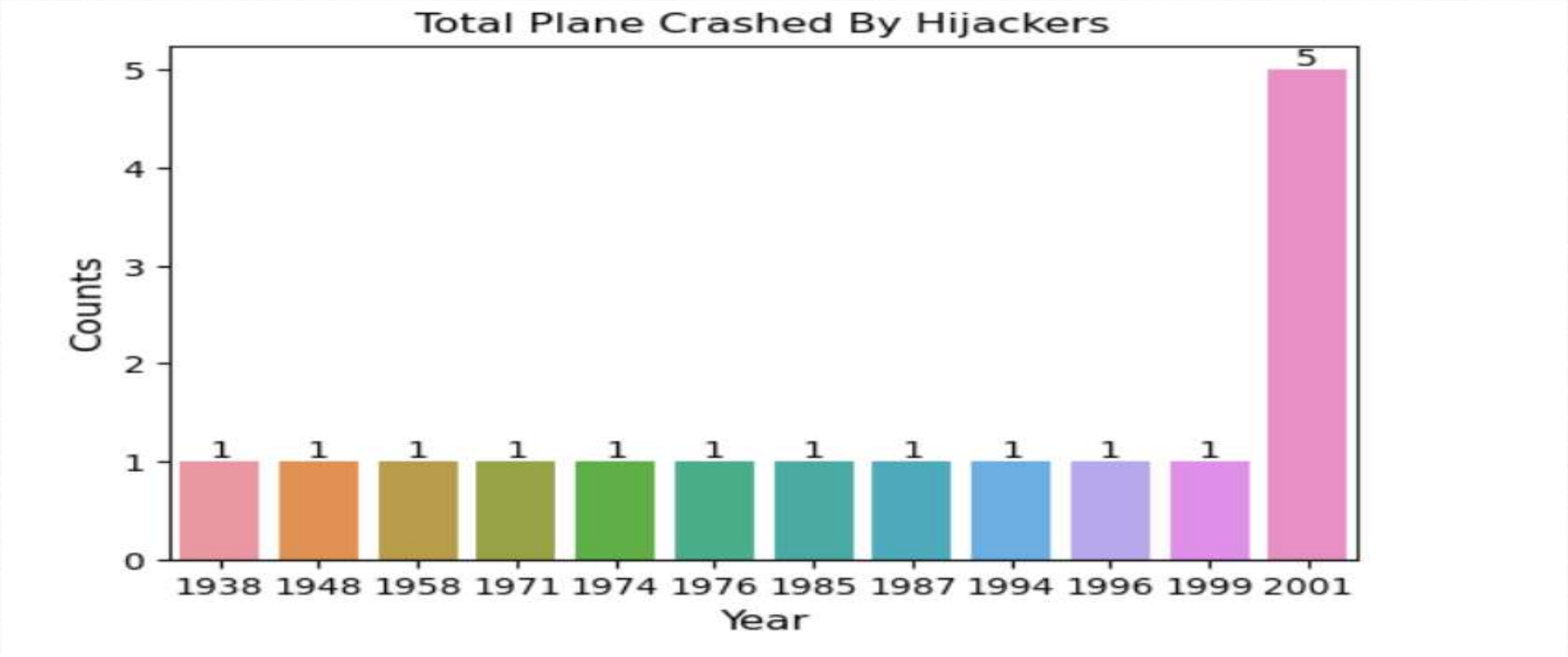


Conclusions:

1- From the above figure, we can see that the Douglas DC-3 type of plane had 331 crashes.

2- As we can see in the graph above, the Douglas DC-3 had the most fatalities with 4807, followed by the Boeing 767-223ER

Figure - 13



Conclusions:

1- According to the figure above, the highest number of plane crashes occurred in 2001, with 5 incidents.

Figure-14

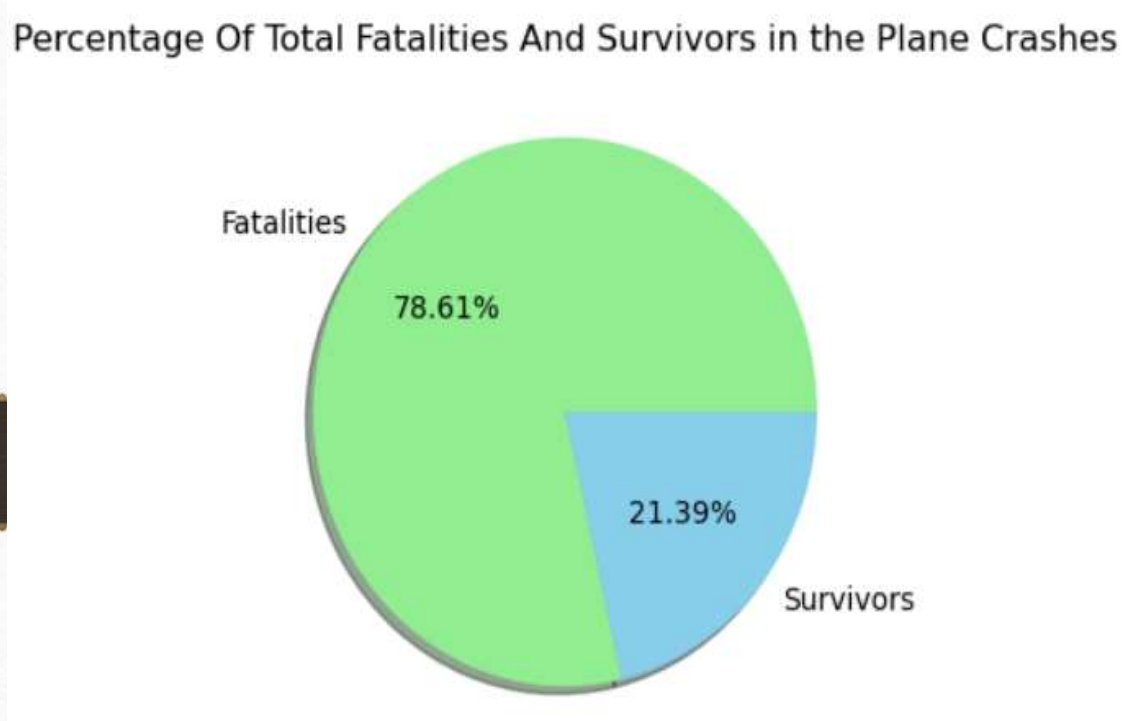
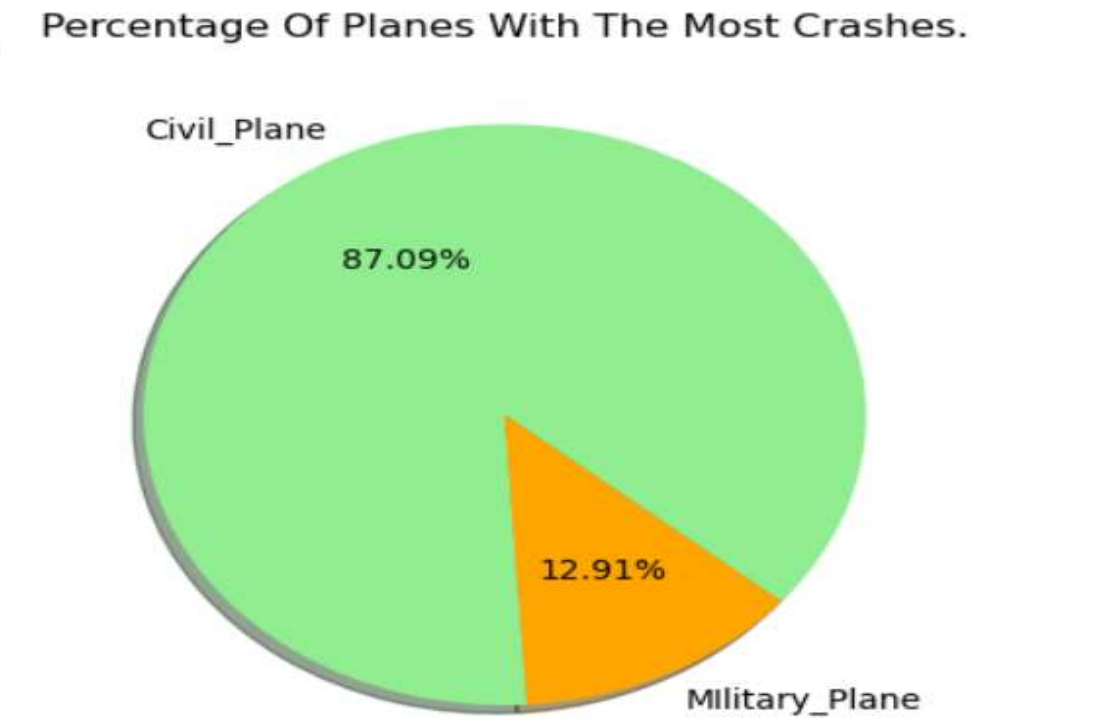


Figure-15



Conclusions:

1-As we can see in the above **figure- 1**, only 21.39%(30752 in numbers) of people survived in plane crash accidents, while 78.61%(113039 in numbers) died.

2- In the **figure-2**, 87.09% of the crashes were civil planes, while 12.91% were military planes.

Figure-16

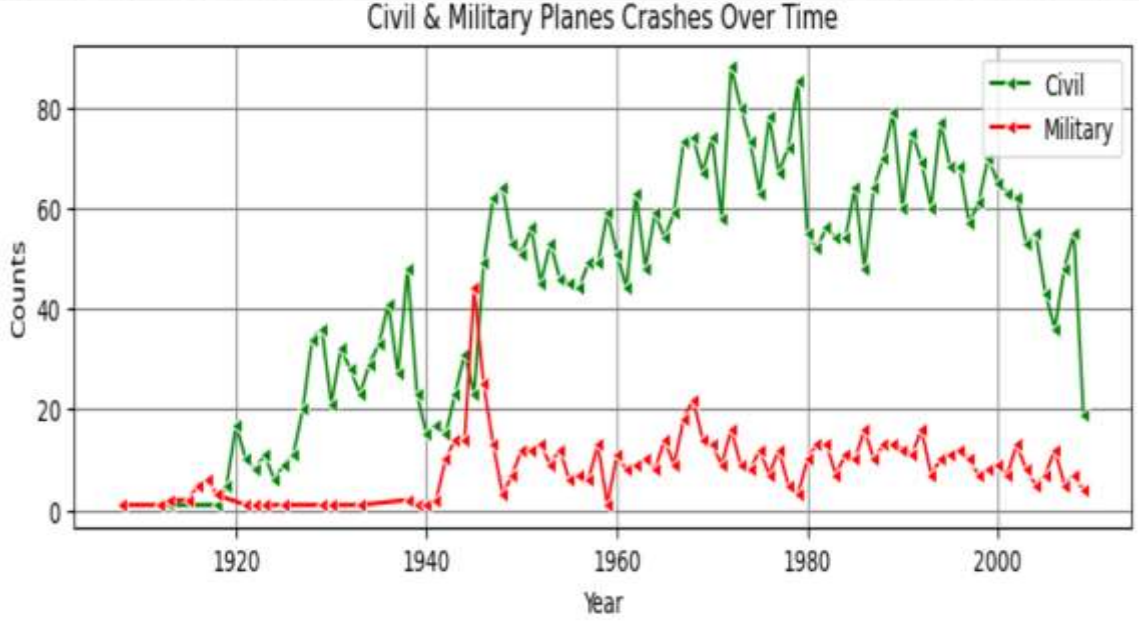
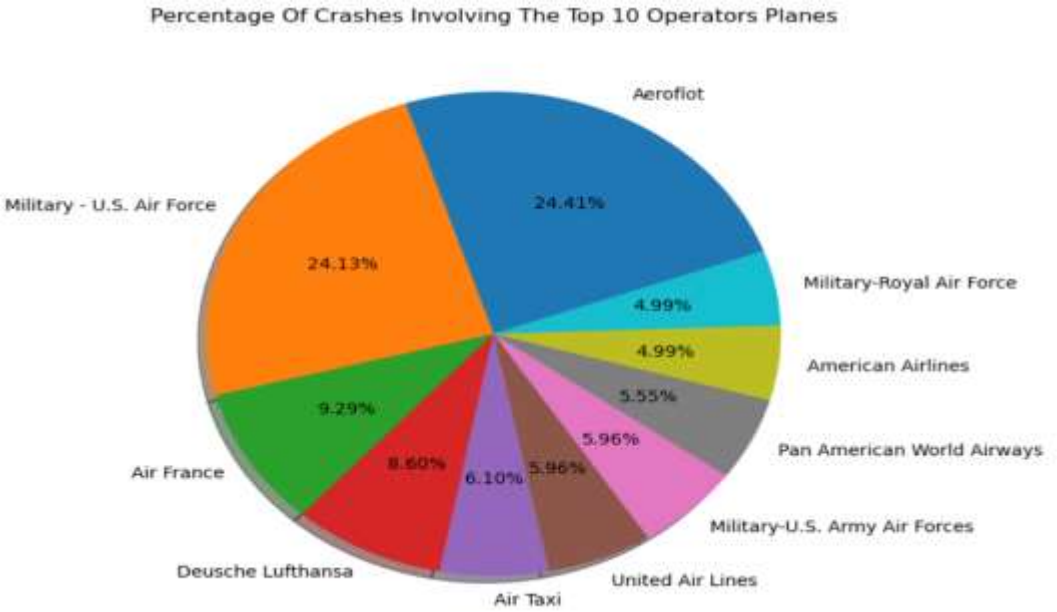


Figure - 17



Conclusions:

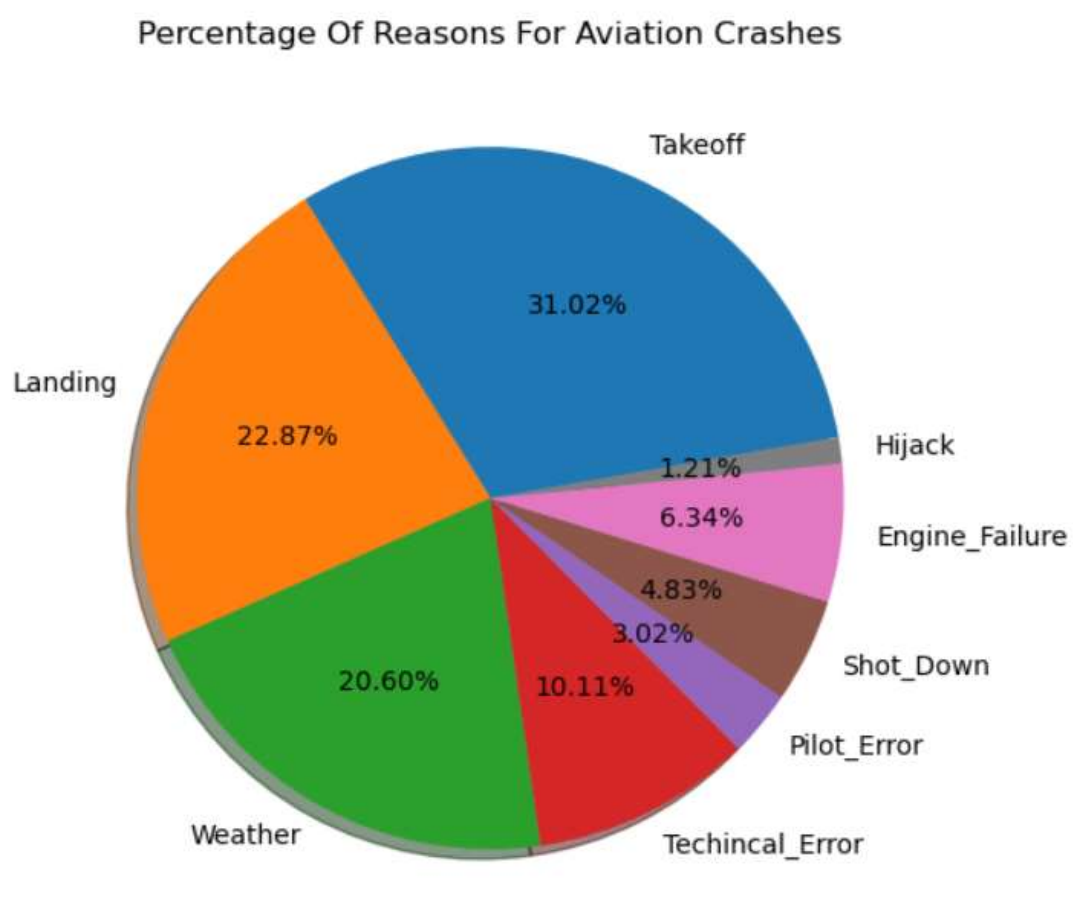
1- We can see in the figure above that the highest number of military plane crashes occurred in 1942, with around 44 incidents.

2- The maximum number of Civil Plane Crashes in 1972 was around 85 incidents.

3- Compared to military plane crashes, civil plane crashes occurred more frequently.

As we can see, Aeroflot has the highest percentage of crashed planes at 24.41%, followed by the U.S. Air Force has the second highest percentage at 24.12%.

Figure -18



Conclusions:

As we can see in this figure, the reasons of the plane crashes are "Take off, Landing , Weather, Technical_Error , Pilot's mistake, Shot_Down, Enginge Failjure , Hijack"

Concluison:-

The analysis of plane crash data from 1908 to 2008 reveals several critical insights into aviation safety trends and areas needing improvement. The years 1972 and 1968 experienced the highest number of crashes, with 104 and 96 incidents, respectively. Additionally, 1972 recorded the highest number of fatalities, with 2,937 deaths. Fatality rates peaked in 1972 and 1985, with 2,937 and 2,670 fatalities, respectively, reflecting a period of increased risk as passenger numbers grew between 1960 and 2000. Aeroflot and the U.S. Air Force reported the highest crash rates, with 176 and 174 incidents, respectively, and Aeroflot also had the highest fatality count at 7,044, highlighting serious safety concerns. The Douglas DC-3 was notably involved in the most crashes, with 331 incidents and 4,807 fatalities, indicating specific issues with this aircraft model. Training routes were particularly hazardous, accounting for 176 crashes and 940 fatalities, underscoring the need for improved pilot training and safety protocols. Survivors comprised only 21.39% of crash victims, while fatalities accounted for 78.61%, emphasizing the severe nature of these incidents. The year 2001 saw the highest number of hijacking incidents, with 5 occurrences, signaling a critical need for enhanced security measures. Civilian aircraft were involved in 87.09% of crashes, compared to military aircraft at 12.91%, suggesting a continued focus on improving safety in the civilian sector. Overall, while there has been a noticeable decline in crashes since 2000, reflecting advancements in safety technology and practices, the data highlights ongoing challenges and the necessity for continued improvements in pilot training, aircraft technology, security measures and should avoid dangerous route and location where most of the incidents occurred.